A Matter of Semantics?

Intelligence, Open Data and the Future of ERM

Jonathan Blackburn

Mason Hall

Mason Hall

Electronic Resource Integration Management Librarian

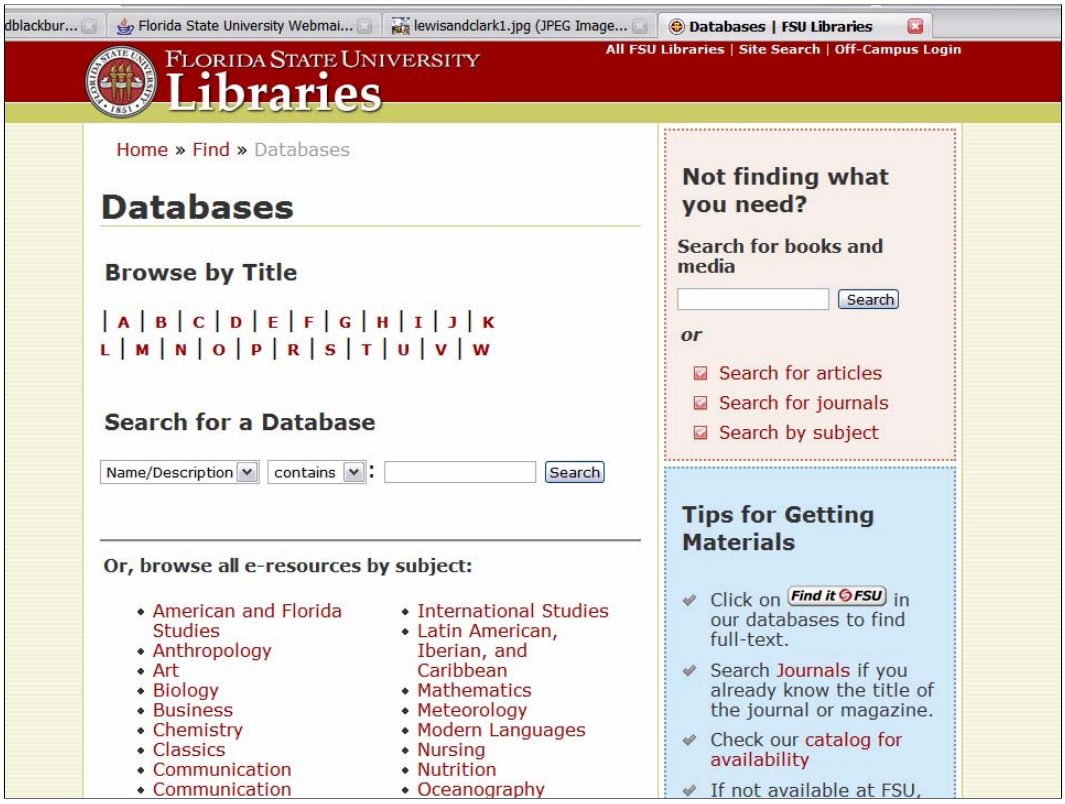
Jonathan Blackburn

Web Development Librarian

So, who are we?

Before we really begin, let me tell you a story that will hopefully introduce the idea behind our presentation.

Several months ago . . .



Jon and Mason were contemplating recreating their database A-Z list from the ground up.

(Not just how it looked, but how each database was described . . .)

The “Perfect” DB Record

- Title
- Alternate Titles
- Description
- Full-text coverage
- Content Types Included (journals, newspapers, etc.)
- Features (RSS, Alerts, Refworks, etc.)

But this takes work to maintain (update) . . .

E.G. Each database has a name, brief description, but also might include full-text coverage dates, publisher, feature list (Refworks? Alerts?), etc.

However …

Who would do the work of maintaining this information?

Information that was duplicated in places like our ERMS, OpenURL Resolver, and Federated Search System.

Those systems, however, did not have features we wanted in our Database List.

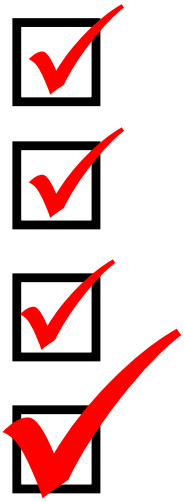
Who?

Librarians at FSU

Librarians Outside FSU

Publishers and Vendors

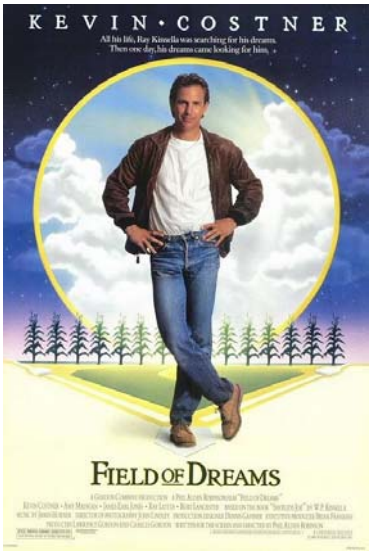
All of the Above



Now, that we had all these OTHER people maintaining this system, what information would we decide to include?

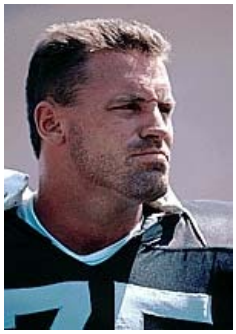
We also asked ourselves the question . . .

Why?



Why hasn’t anyone built a system like this already?

How?



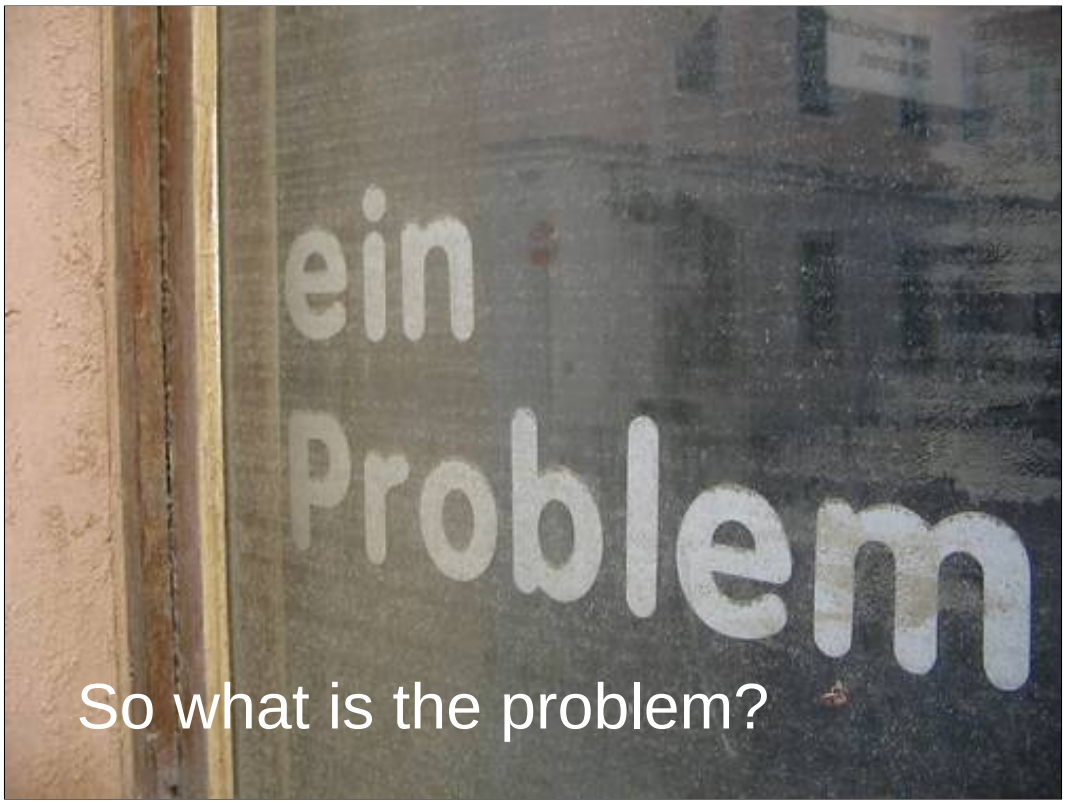
How we might build a system like this, if possible?

So, began our quest . . .

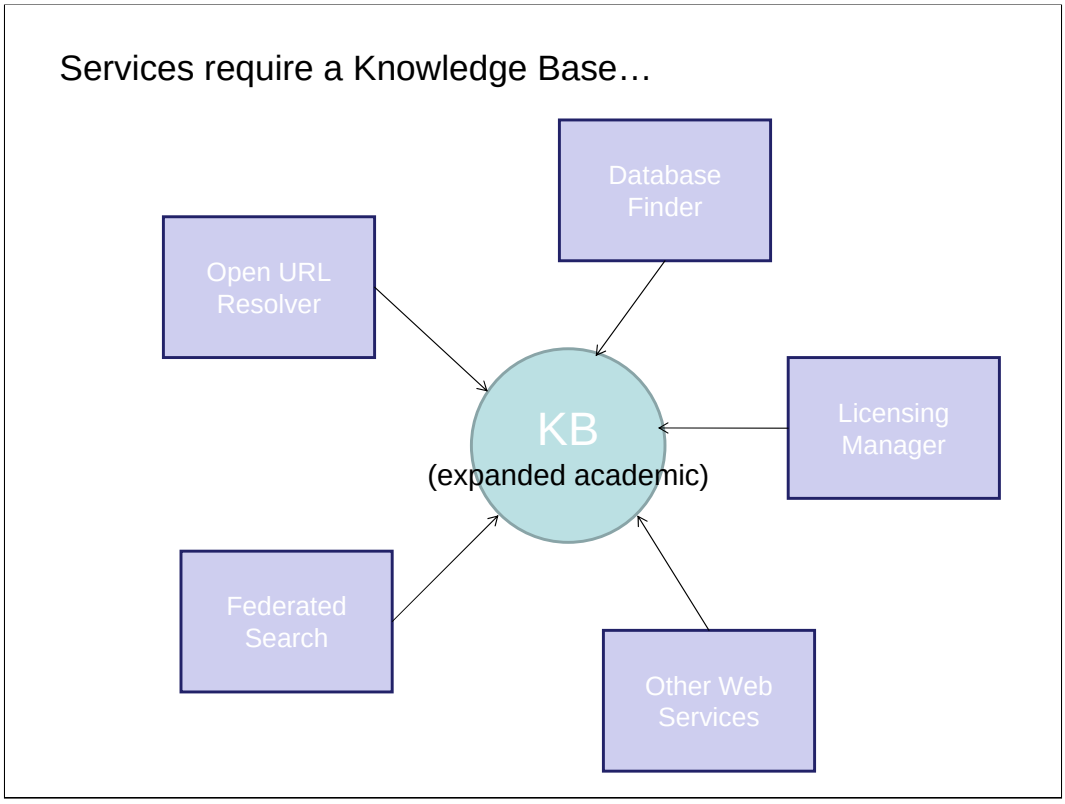
To answer these questions and make sense of this whole issue.

We don’t claim to have answers to all these questions, but I do think we have a pretty decent grasp of the problem now.

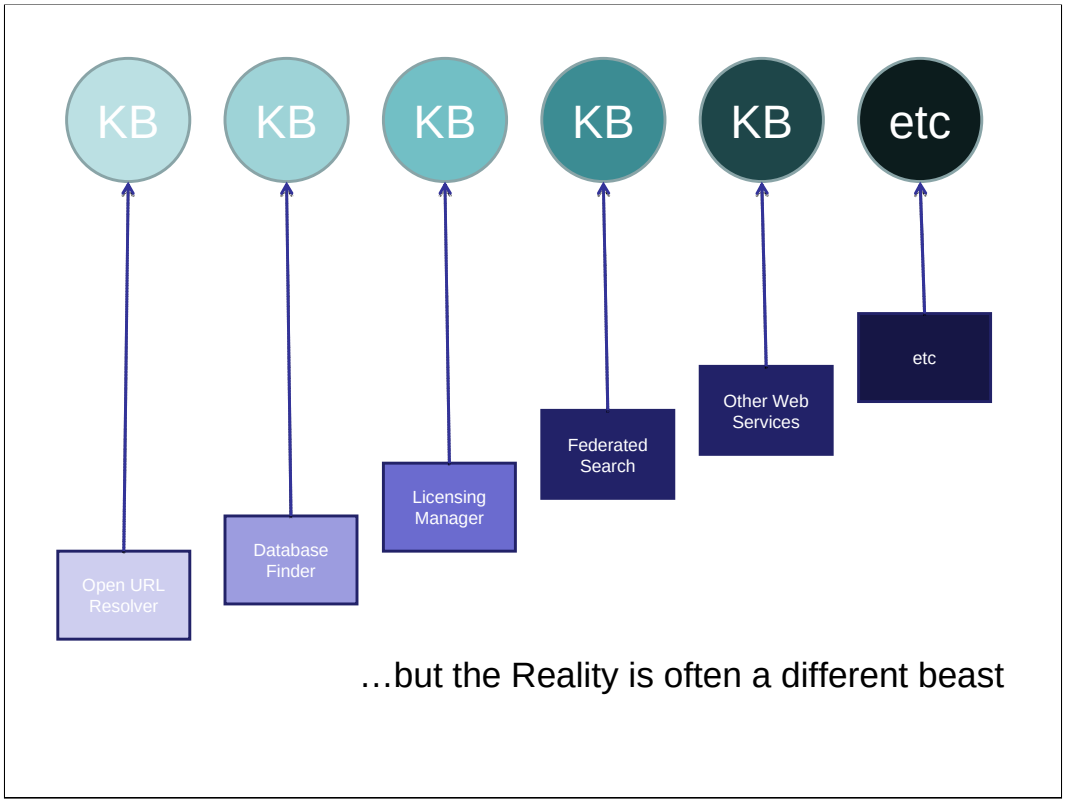
So, w/ no further ado . . . I will hand things over to Mason.


ein
Problem
So what is the problem?

Duplication of Data



One place to put everything … where services can find this information.

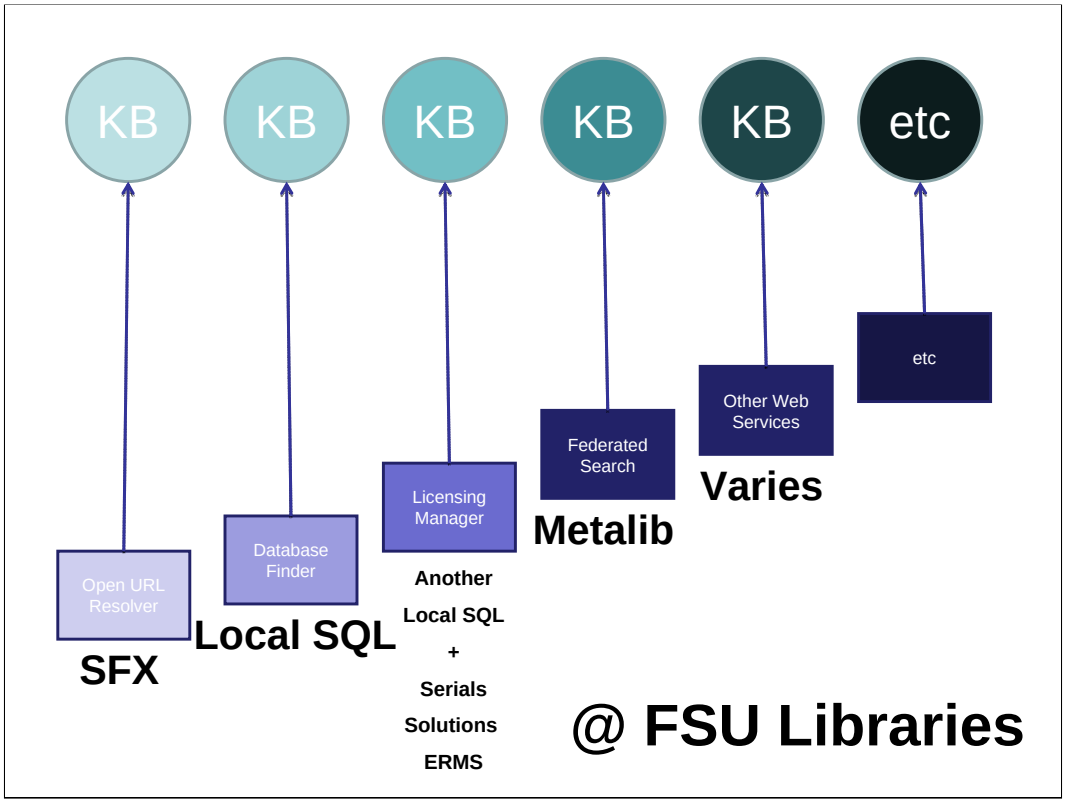


Each service has its own KB…..

Give example of single institution

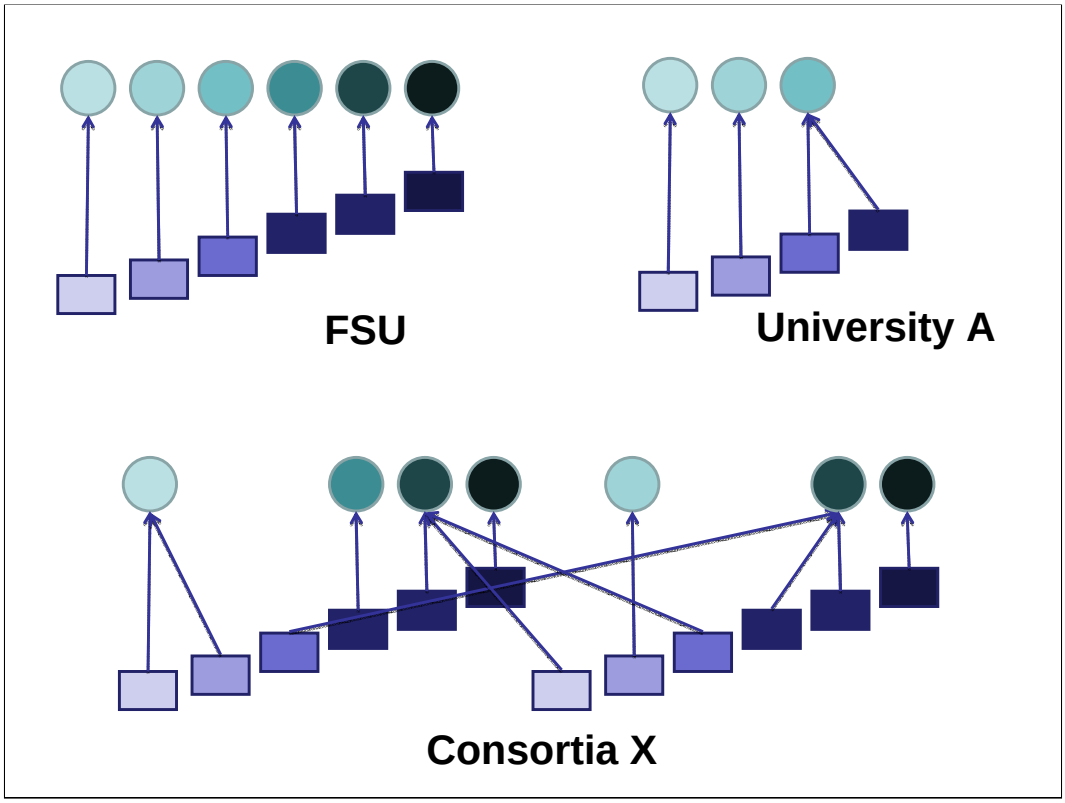
TODO:

Add notes, if possible, next to each “service” (e.g. SFX, etc.)



CHANGE:: increasing # of institutions … show that each institution may use several knowledgebases in combination…

May share, but overall, there is great duplication of data.

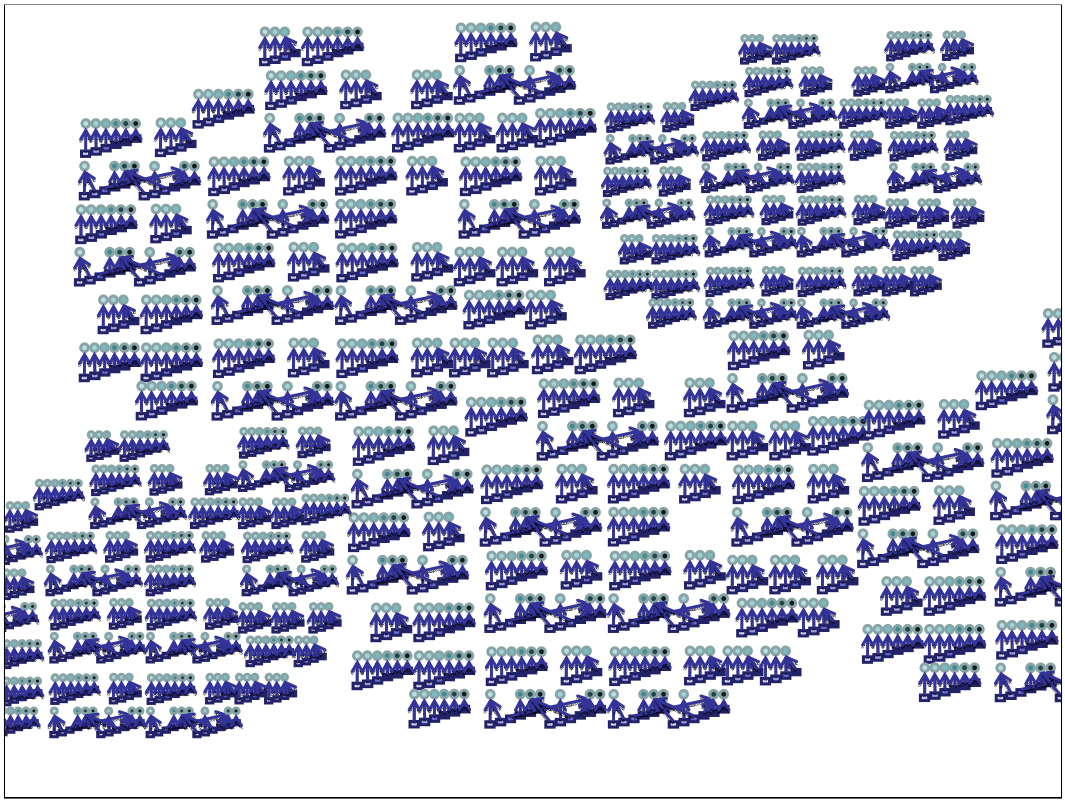


Each service has its own KB…..

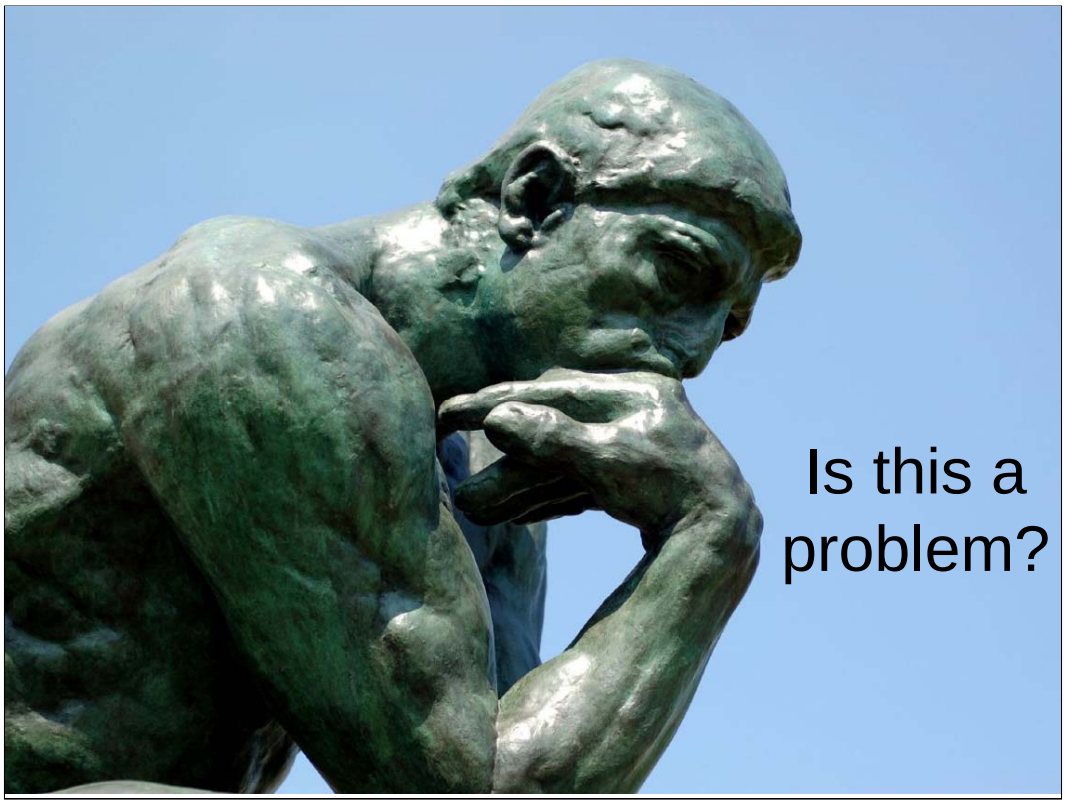
Give example of single institution

TODO:

Add notes, if possible, next to each “service” (e.g. SFX, etc.)



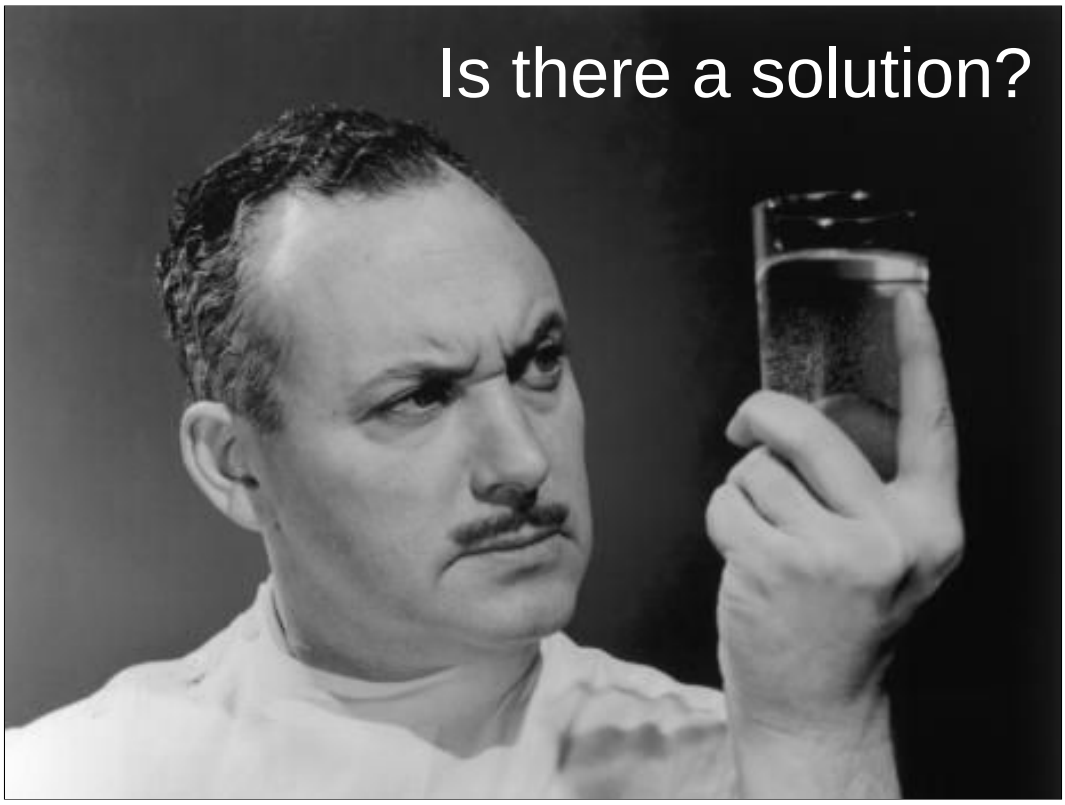
The idea that data is currently owned and updated by vendors



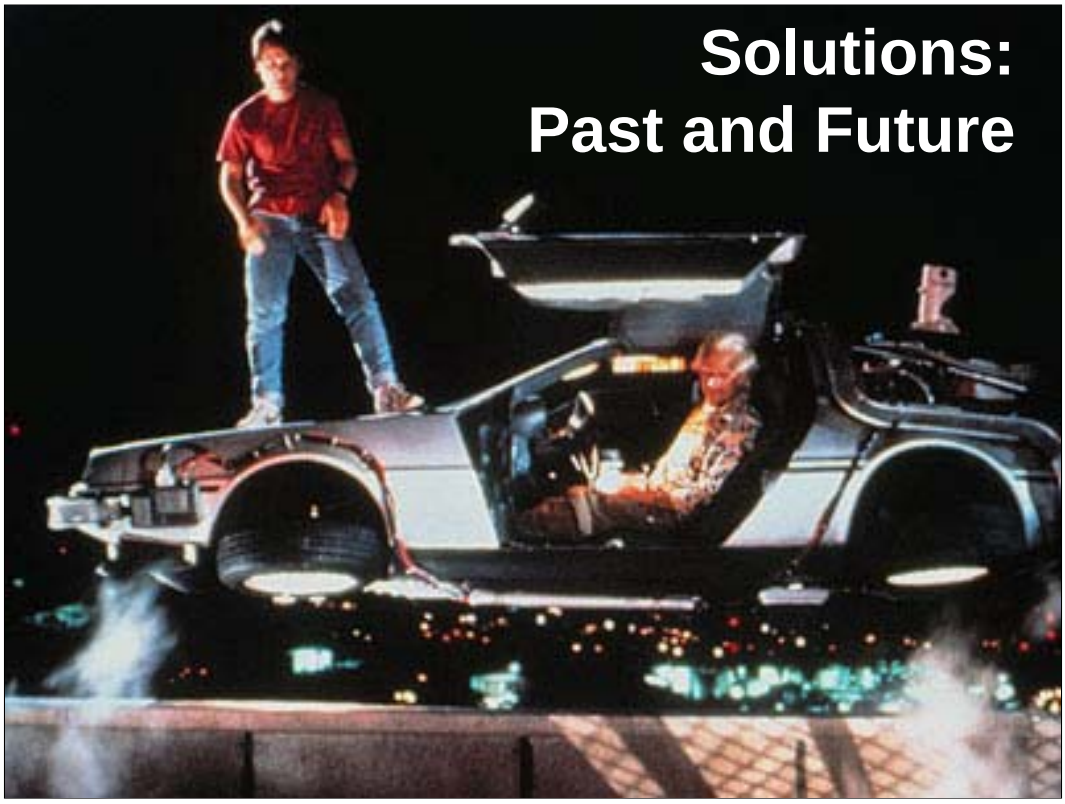
Some ways it’s not … lot of work!

Reliant on them for updating, accuracy, format (use??), and robustness of information provided. (They choose fields/attributes to describe various journals, databases, etc.)

Wouldn’t be great if we could --- if YOU could???



So, now we know what the problem is . . . Can we find a solution?



We are not the first ones to ask this question . . . Others much smarter than us have lead projects devoted to addressing aspects of this problem.

Jake

- Formerly at: http://jake.openly.com/

jointly administered knowledge environment

CUFTS

http://researcher.sfu.ca/cufts

More recent project…

(show interface)

-Tied to their opensource CUFTS OpenURL service
-Maintained by SFU staff but exploring “collaborative model”
-Used to “sell” subscriptions to KB, but now free (business model??)

Ockham (main one I am going to talk about)

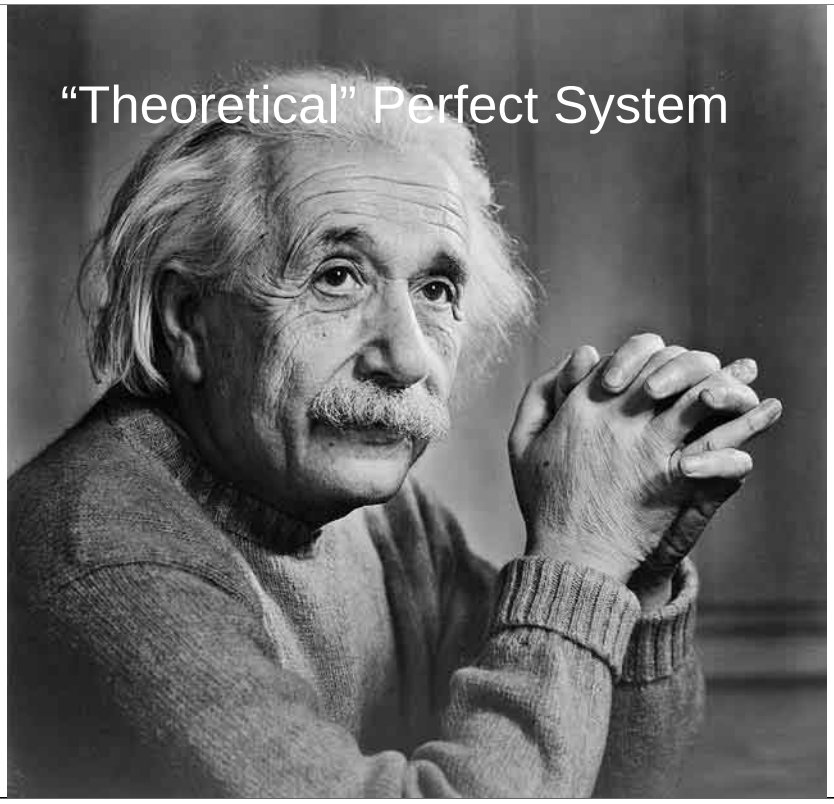
Headed up by a # of libraries, but I spoke recently to Jeremy Frumkin…

Idea taken from British project (IESR) …

Way for machines and humans to “discover” new services ….

So, currently, nothing I can show you… but the way it would work is ….

“Theoretical” Perfect System




Library OKRA?

Library OKRA

"**One Knowledgebase to Rule Them All**

One Knowledgebase to Find Them

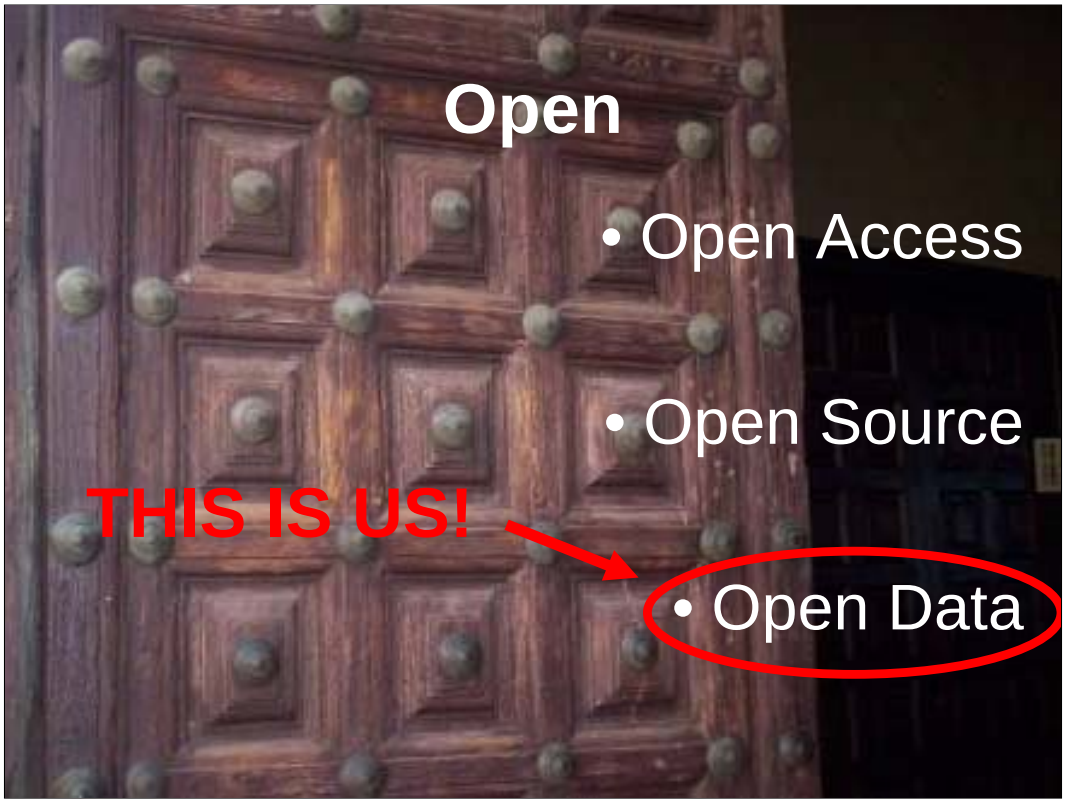
One Knowledgebase to Bring Them All

And In the Darkness Bind Them "

- The Lord of the … ERM?

JON>>

NOW THAT WE HAVE a fancy ACRONYM for it, what might this hypothetical system look like?



We CAN re-use…

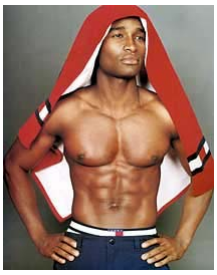
We can ensure accuracy…

We can ensure relevance (of metadata)…

Institutions can choose to host locally customized instances of knowledgebase.

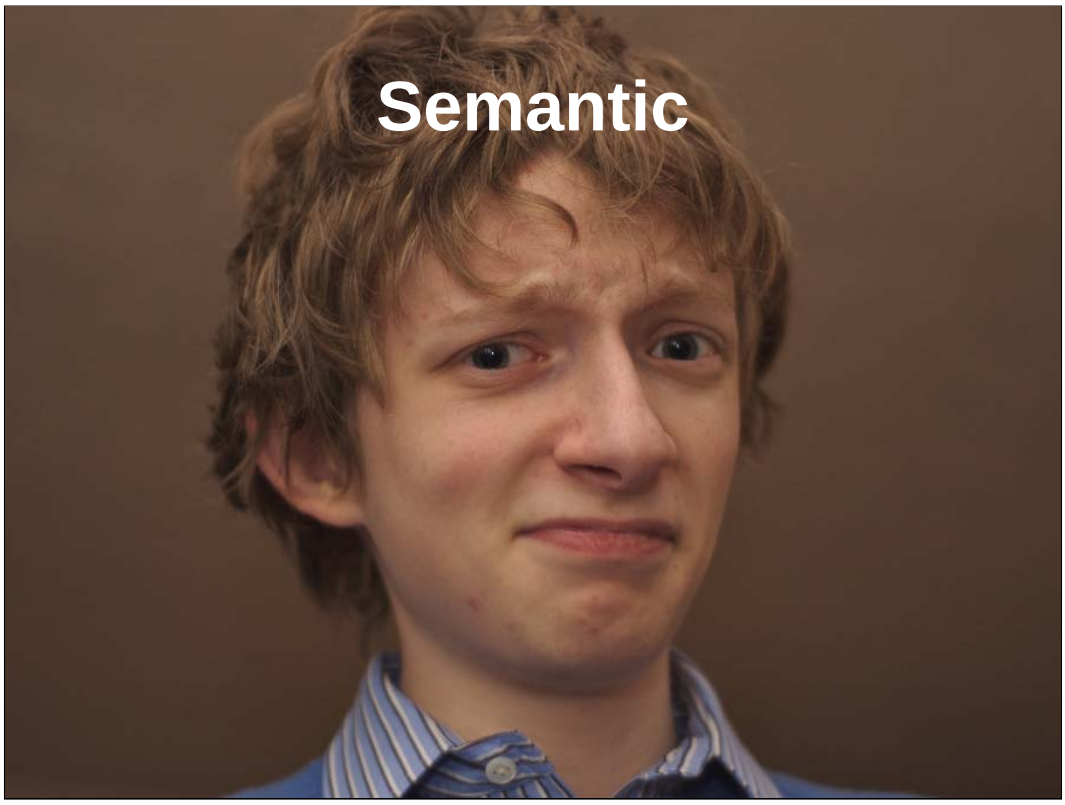
Also, openness is key to getting people to contribute to the KB, which is the next thing I would like to discuss…


Shared



Models
(For Sharing)

- Cooperative Cataloging
- Wikipedia
- Freebase.com (the website)



Machine readable

Define semantic web with examples – rdf, owl, etc then with actual use

Relationships between entities

Why this is useful

Semantic = Relationships

A very brief lesson

Subject → Predicate → Object

Publisher *publishes* Journal
Vendor *owns* Database
Article *has* author(s)

Triples

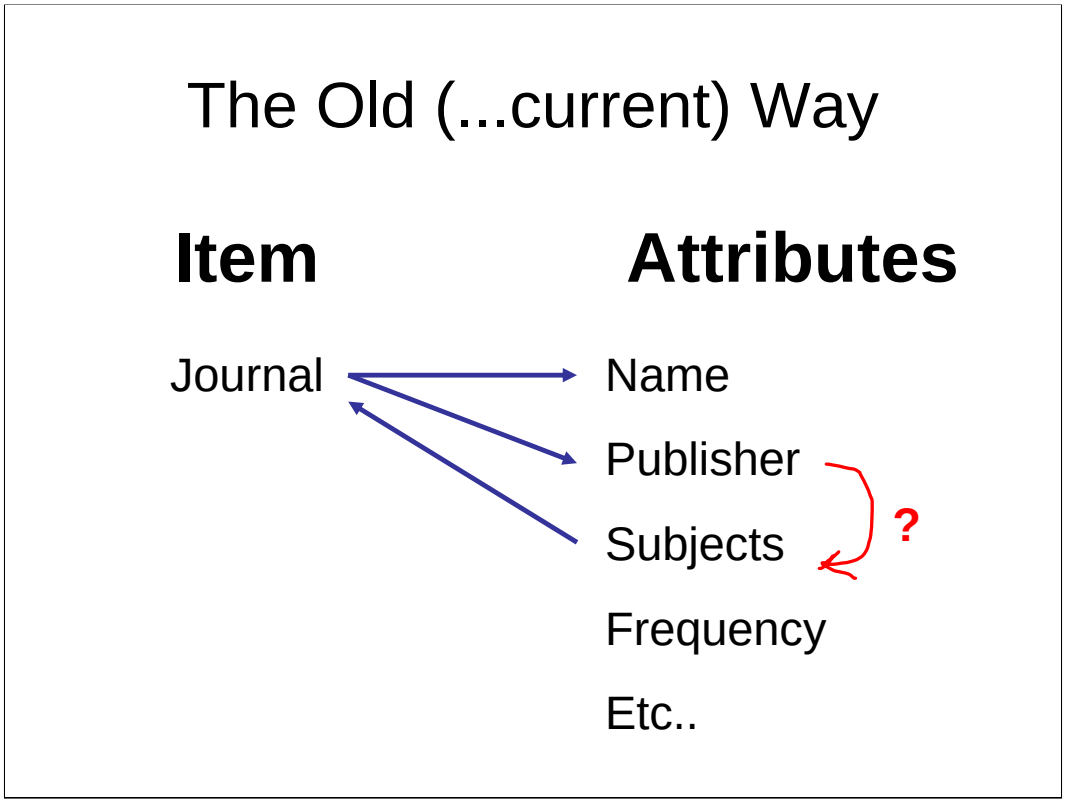
Subject – Predicate -- Object Slide (Relationships between objects)

“triples”

Publisher –owns – Journal
Database – provides – journal coverage

Company >> provides >> Interface
Interface >> includes >> Database
Database >> includes >> Journal XYZ/Coverage Date
= Company >> provides interface >> Journal XYZ/Coverage Date
Coverage Date >> Which Journals

Metadata can be incrementally improved … not based on individual records, set metadata scheme – but flexible, extensible, etc.

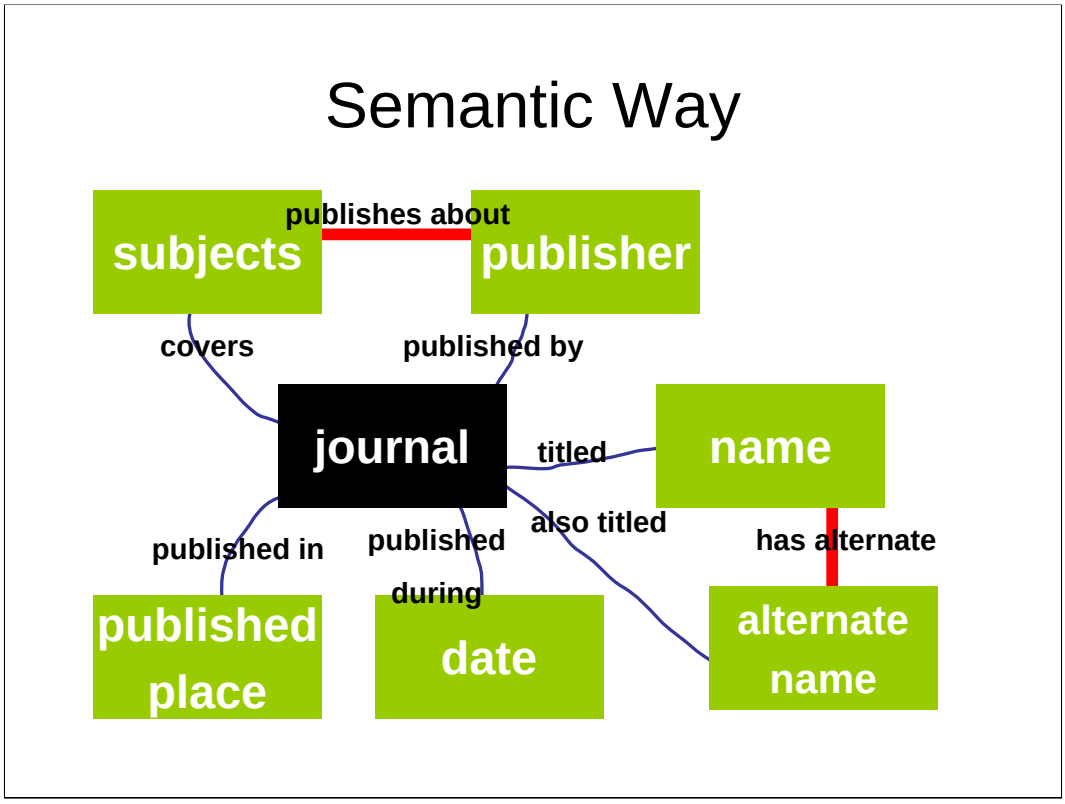


Catalog Record

Title
Attributes
.. Author
.. Publisher
.. Subjects

Titles on a subject, by an author

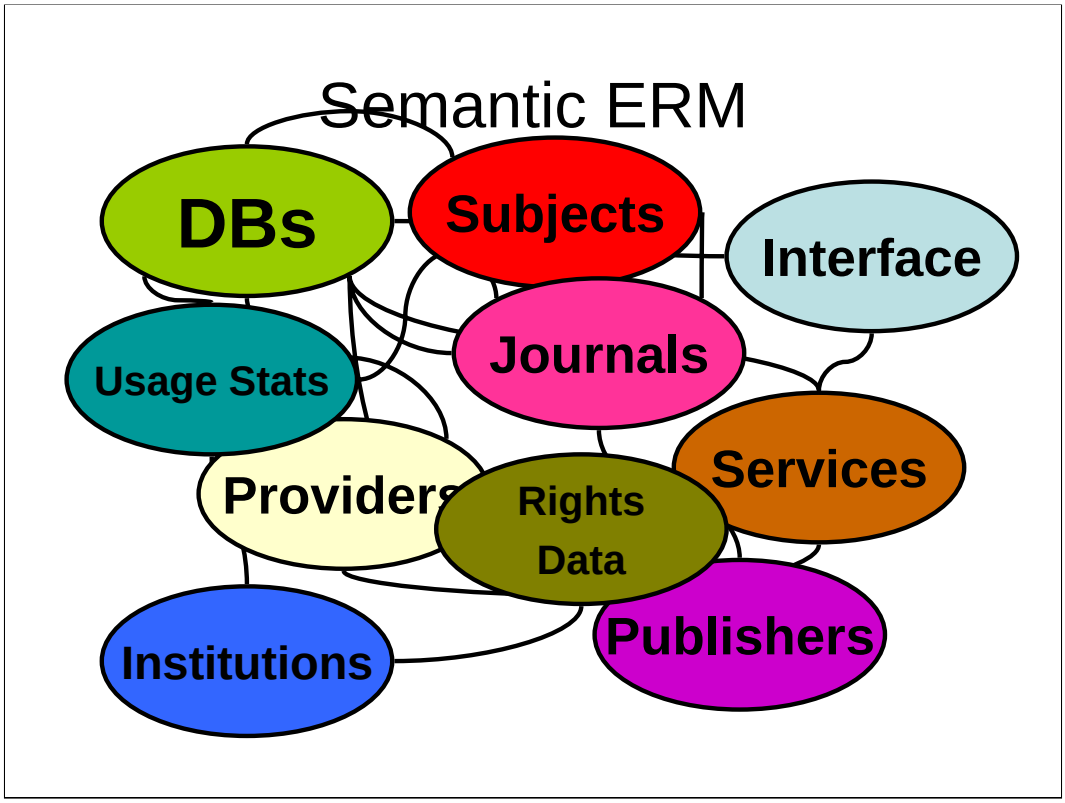
Subjects written by a particular
Authors who write on a particular subject
Publishers publish on a particular



Author writes Book

Book has Publisher

Book has Subjects


Semantic ERM
DBs
Subjects
Interface
Usage Stats
Journals
Services
Providers
Rights
Data
Publishers
Institutions

What are you describing? (Service oriented model of description… object themselves… reusable for multiple services)

Institution-specific data

How strict do you want the metadata? (Even semantic stuff has standards… RDF, DC, etc.?)

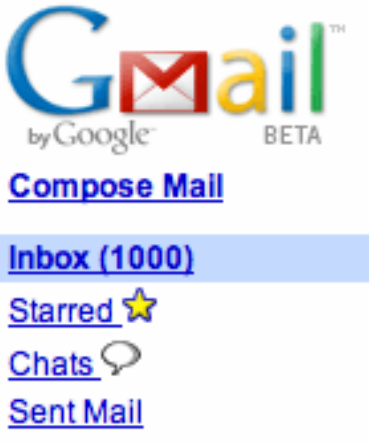
Incentives?? (Why would an instituion volunteer to contribute time and resources to maintaining this KB?)

The consensus (from the people we talked to)…

The “Killer App”

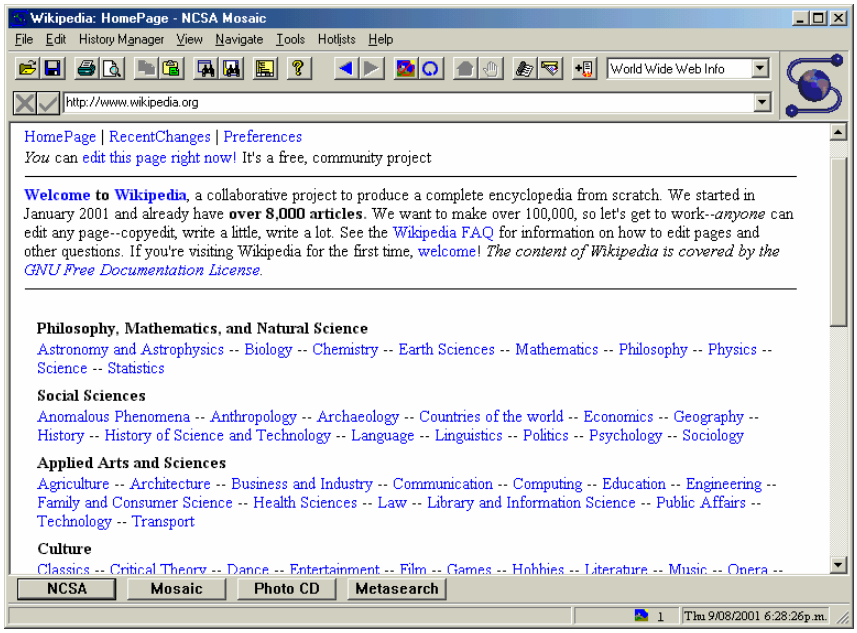
Other “Killer Apps”

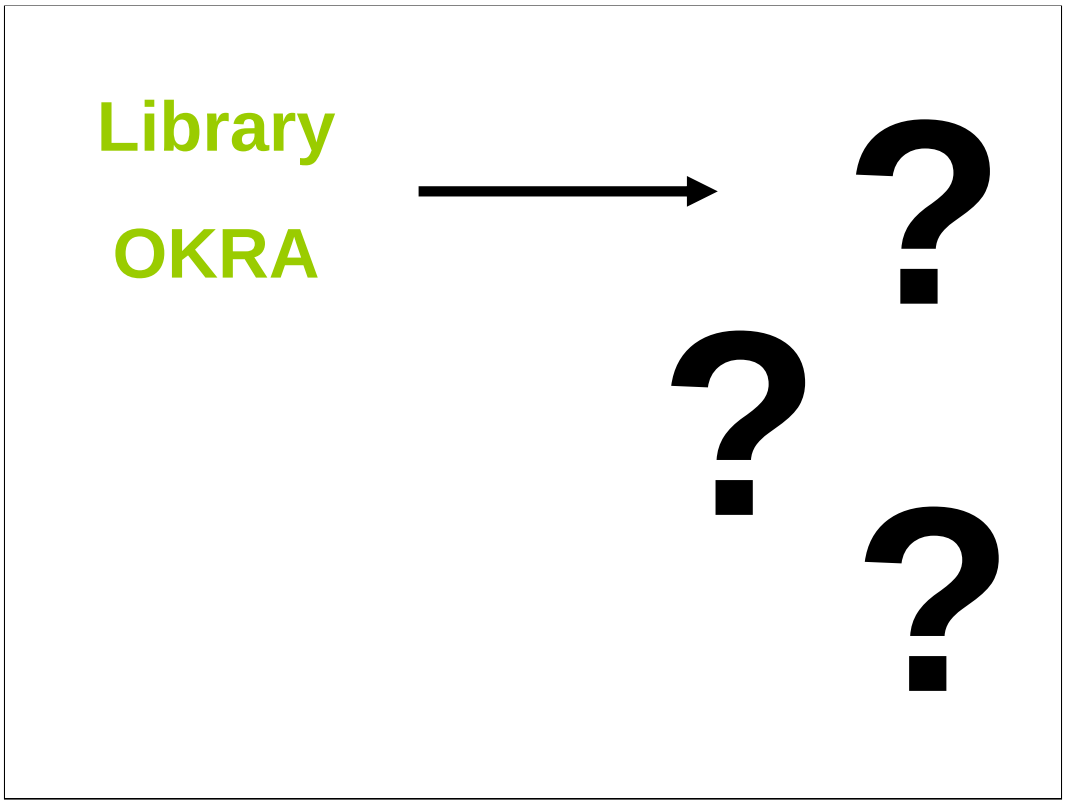
Internet → E-mail



What is a “killer app”?

WWW → Mosaic


Wikipedia: HomePage - NCSA Mosaic
File Edit History Manager View Navigate Tools Hotlists Help
World Wide Web Info
http://www.wikipedia.org
HomePage | RecentChanges | Preferences
You can edit this page right now! It's a free, community project
Welcome to Wikipedia, a collaborative project to produce a complete encyclopedia from scratch. We started in January 2001 and already have over 8,000 articles. We want to make over 100,000, so let's get to work--anyone can edit any page--copyedit, write a little, write a lot. See the Wikipedia FAQ for information on how to edit pages and other questions. If you're visiting Wikipedia for the first time, welcome! The content of Wikipedia is covered by the GNU Free Documentation License.
Philosophy, Mathematics, and Natural Science
Astronomy and Astrophysics -- Biology -- Chemistry -- Earth Sciences -- Mathematics -- Philosophy -- Physics -- Science -- Statistics
Social Sciences
Anomalous Phenomena -- Anthropology -- Archaeology -- Countries of the world -- Economics -- Geography -- History -- History of Science and Technology -- Language -- Linguistics -- Politics -- Psychology -- Sociology
Applied Arts and Sciences
Agriculture -- Architecture -- Business and Industry -- Communication -- Computing -- Education -- Engineering -- Family and Consumer Science -- Health Sciences -- Law -- Library and Information Science -- Public Affairs -- Technology -- Transport
Culture
Classics -- Critical Theory -- Dance -- Entertainment -- Film -- Games -- Hobbies -- Literature -- Music -- Opera --
NCSA
Mosaic
Photo CD
Metasearch
Thu 9/08/2001 6:28:26p.m.


Library
OKRA
?
?
?

Is there hope?

- Encourage Publishers to be more Open with Data
- Press Vendors to Uncouple Data and Services
- Support Open Source Library Software

Encourage publishers and library vendors to be more “open” w/ current data (and how its licensed or provided to libraries).

Support opensource library software software and solutions… like CUFTS? (question: will os ILS’s like Koha and Evergreen have a ErM component … if so, they will need a KB.)

Don’t be satisfied w/ current opensource and commercial solutions … search for the killer app!

What are the current systems NOT doing?? If there is something, this might be a feature that LIBRARY OKRA (and its related apps) can exploit?

Continue this Discussion on:

LibraryOKRA.com